Viral transduction and the dynamics of bacterial adaptation

Philippe Cherabier[1*], Sylvie Méléard[2], and Régis Ferrière[1,3,4]

[1]Institut de Biologie de l'École Normale Supérieure (IBENS), Université Paris Sciences et Lettres, CNRS, INSERM, 75005 Paris, France

[2]CMAP, CNRS, École polytechnique, Institut polytechnique de Paris F-91128 Palaiseau, France et Institut Universitaire de France.

[3]Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

[4]International Research Laboratory for Interdisciplinary Global Environmental Studies (iGLOBES), CNRS, ENS-PSL University, University of Arizona, Tucson, AZ 85721, USA

[*]Corresponding author: philippe.cherabier@bio.ens.psl.eu

Abstract

Viral infections can exert a large influence on their hosts' ecology by causing widespread mortality, but they also shape the evolutionary adaptation of hosts in a number of ways. A major pathway for viruses to do so is through the transfer of genetic material among individual hosts, a process known as transduction. While horizontal gene transfer is known as a major factor in prokaryotic macroevolution, its role in the microevolutionary adaptation of hosts populations is poorly known. By facilitating the transfer of beneficial alleles between host cells, transduction might facilitate and accelerate bacterial adaptation. Conversely, the risk of transferring deleterious alleles may hinder and slow it down. Here we resolve the effect of transduction on bacterial adaptation in a simple eco-evolutionary model for the combined dynamics of transduction and adaptive evolution of an ecological (resource-use) trait. The transfer of beneficial alleles by tranduction speeds up adaptation whereas the transfer of deleterious alleles causes strong stochastic fluctuations of the trait value around the adapted value. In contrast to the expected effect of recombination, which tends to oppose phenotypic diversification, viral transduction can increase host phenotypic diversity.

1 Introduction

Viruses are global actors of Earth's ecosystems. In the oceans, they are the most abundant life form, overnumbering bacteria by an order of magnitude on average (Wommack and Colwell 2000). Their contribution to prokaryotic mortality is estimated to be in the range of 10%-30%, with as much as a 100% mortality for some species (Danovaro et al. 2011, Wommack and Colwell 2000). With the more abundant bacterial species being more at risk of viral infection, viruses can increase biodiversity through a *kill the winner* mechanism (Thingstad and Lignell 1997), and bacterial lysis is a significant source of nutrient for primary producers (Van Hannen et al. 1999). The direct and indirect functional effects of viral infection and mortality (Breitbart et al. 2018) extend across all levels of organization, from single-cell organisms to the megafauna (Munn 2006).

But the impact of viruses in shaping ecosystems goes beyond the induced mortality. Indeed, through transduction, a phage-mediated horizontal gene transfer mechanism, viruses can also influence the genotypic composition and diversity of their host population, hence potentially the dynamics of evolution.

Horizontal gene transfer is thought to be a major contributor to evolution (De la Cruz and Davies 2000, Ochman et al. 2000, Lawrence 2002). Contrary to vertical gene transfer, where genes are passed on from parent to offspring, a horizontal gene transfer event sees genes passed on from neighbor to neighbor. It allows for a rapid spread of innovation (Hall et al. 2017, Hao and Golding 2006), but it can also spread deleterious genes (Redfield 1988, Billiard et al. 2016). There are three big types of mechanisms of horizontal gene transfer: transformation, conjugation, and virus-mediated transfer or transduction. Transformation consists in cells picking up free-floating DNA and integrating it to their own genome; conjugation consists in plasmid-mediated transfer between two cells; transduction consists in bacteriophage-like particles transfering DNA between cells.

Studies have long been emphasized on the macroscopic role of horizontal gene transfer in evolution, most notably through the wide-spread of innovation: for example, we believe that many innovations that appear in gut bacteria are the result of horizontal gene transfer (Foster et al. 2017), and that entire biological pathways in oceanic eukaryotes are the result of transduction (Monier et al. 2009). But incremental evolutions studied in adaptive dynamics

are rarely studied through the lense of horizontal gene transfer (Raz and Tannenbaum 2010), and transduction in particular needs to be more closely investigated (Touchon et al. 2017).

The conventional hypothesis is that by facilitating the transfer of genetic elements between host cells, GTAs facilitate and accelerate the adaptation of their bacterial host population to a changing environment (Andam et al. 2015). But if recipient bacteria can rapidly acquire advantageous mutations through transduction, so can they receive deleterious ones. In this case of evolutionary trade-off, in which conditions does transduction help or hinder adaptation?

Our study offers a novel modelling of transduction events in an attempt to study the impact such a mechanism can have on bacterial adaptation. We start with an individual-centered model that we rescale for large population and at ecological, then evolutionary timescales. We find that transduction can change the evolutionary trajectory of the host by broadening the spectrum of potentially invading mutants, adding a stochastic term to the canonical equation. We also find that small transduction rates cause an acceleration of adaptation.

2 Methods

2.1 Individual-level model of infection and transduction

Our model of transduction is summarized in figure 1a. During the reproduction and encapsulating of viral genome, an error may occur resulting in a viral particle containing part of the host genome instead of viral DNA (general transduction, Touchon et al. 2017). These viral particles, called *Genetic Transfer Agents*, or GTA, are no longer able to cause the lysis of the cells they infect. Instead, they pass the genetic material from the previous host on to the receiving cell, where it may be integrated in the genome by recombination. As a consequence, the phenotype of the receiving cell, and its offspring lineage, may be altered (Fig. 1a).

To describe the genotypic and phenotypic effects of transduction, we focus on a quantitative character, or trait, a, and introduce a simple model of the genotype-phenotype map. We make the following assumptions:

1. The trait is under the control of many loci of small, additive effects. The effect of mutation is also treated additively.

2. We use the principle of an infinite site model, so that no two mutation will impact the same locus.

3. Alleles at the different loci can be moved by transduction. When a transduced enters a bacterium's genome, it is integrated through *non-homologous end joining* (Popa et al. 2011), i.e. the new DNA fragment is *added* to the host genome, and does not replace the resident allele.

4. We assume no dosage effect, meaning that the number of occurrences of an allele is irrelevant to the intensity of expression of the gene.

5. The "packaging error" of host DNA in a new viral particle (GTA), the delivery of the previous host's DNA by the GTA, and the integration of transferred genetic material into the newly infected host's genome, are stochastic events. Each occurs with a certain probability.

6. The integration step of the transduction process is more likely if the phenotypes are more similar. The reason is that the underlying genotypic difference between more similar phenotypes is likely smaller, which may facilitate recombination.

As described by assumptions 3 and 4, a transduction event does not overhaul the genotype of the recipient bacteria, but only adds to it. This means that when cells receive a copy of a mutation they already carry, two occurrences of the same mutation will be found in their genotypes, and yet, per assumption 4, their phenotypes will not be altered. Only the presence or absence of an allele influences the cell's phenotype, leading to the following model of a genotype-phenotype map.

Let n be the number of mutations that have occurred in a population at time t. Since each mutation occurs on a different site (assumption 1), we can define $\mathcal{M}_i \subseteq \{1, ..., n\}$ the subset of mutations that individual i carries. We can then measure the effect of each mutation $m \in \{1, ..., n\}$ as a small perturbation ϵ_m, such as the trait a_i for individual i is

$$a_i = a_0 + \sum_{m \in \mathcal{M}_i} \epsilon_m \tag{1}$$

where a_0 is the initial trait of an isogenic population.

Let $i, j \in \mathbb{N}$ be two individuals respectively bearing $\mathcal{M}_i, \mathcal{M}_j$ as mutation sets. Let's consider the case where the individual i meets a GTA originating from individual j, thus bearing $\mathcal{M}_j$. By injecting foreign DNA into the individual, it will 'add' the missing mutations to its genotype without deleting the previous mutations held by i. Thus, i's mutation set will change from $\mathcal{M}_i$ to $\mathcal{M}_i \cup \mathcal{M}_j$, altering its phenotype. It is important to note here that this modelling decision doesn't allow for trait-based transduction alone. Indeed, if $\mathcal{M}_i$ transcribes to a_i and $\mathcal{M}_j$ to a_j, there is no straightforward way to deduce $a_{i \cup j}$ from a_i and a_j alone.

Figure 1b shows how transduction impacts the genetic history of populations. When taking the coalescent tree tracking mutations, we see that a transduction event taking place between two distinct lineages result in the creation of a third lineage comprised of a recombination of the two original genetic material. In that sense, we can see transduction events as sparse sexual reproduction events (Baltrus et al. 2008).

We now consider the simplest case of a resident population in which a mutant appears. Under the modelling described, the resident population has the empty set $\emptyset$ as its mutation set, and the mutant has $\{1\}$ as its mutation set, i.e.:

$$\mathcal{M}_{res} = \emptyset \tag{2}$$

$$\mathcal{M}_{mut} = \{1\}. \tag{3}$$

We conclude that the following transduction events respectively have as consequences:

- Interaction between a bacteria and a GTA of its type: as before, nothing happens.
- Interaction between resident bacteria and mutant GTA: since $\emptyset \cup \{1\} = \{1\}$, the resident bacteria will change its phenotype from resident to mutant phenotype.
- Interaction between mutant bacteria and resident GTA: since $\{1\} \cup \emptyset = \{1\}$, the mutant bacteria will not undergo a phenotype change. Nothing happens.

We see that whether they interact with resident or mutant bacteria, resident GTAs will not alter the phenotype of the receiving bacteria. Transduction in this case is unilateral and the system stays a two-type population at any given time. This also allows us to dismiss the

resident GTA equation in our system, since they do not have any influence on the ecological dynamics.

2.2 Phage-bacteria population dynamics with mutation and transduction

We use an individual based model under the large population limit (Champagnat, Régis Ferrière, et al. 2006). This allows us to analyse bacteria evolution at two different time scales: the ecological time scale of the invasion events, and the evolutionary time scale of the trait substitution sequence (Champagnat 2006). A full mathematical description of the model can be found in the **Supplementary note S1**.

The model describes a population of bacteria (X), viruses (Z) and GTAs (U) with population of order of magnitude K. Bacteria reproduce asexually at intrinsic rate $b(a)$ and die at a rate of $d(a, X)$ which may depend on the population size and trait distribution. We assume a classical multitype logistic model for birth and death rates (Lambert et al. 2005) in order to study the influence of viruses and transduction on adaptation speed and branching. Viruses are considered a homogeneous population, and GTAs are defined by the genotype and phenotype of the original host. Both being viral particles, they die with the intrinsic death rate d_v, and interactions between bacteria and viral particles occur at individual rate $\eta(a)$. An interaction between a bacterium and a virus results in instant lysis and viral burst (V new viral particles are created, each having probability γ to be a GTA). An interaction between a bacterium and a GTA results in the non-homologous integration of foreign genome of trait a' into recipient genome of trait a with probability $\psi(a, a')$, which increases as the difference between a and a' decreases (Touchon et al. 2017).

In the large population limit (large K), the dynamics of the population densities x, y, z, u of, respectively, resident bacteria with trait a, mutant bacteria with trait A, viruses, and GTAs are given by (**Supplementary note S2**):

$$\begin{cases} \frac{dx}{dt} &= (b(a) - d(a) - c(a,a)x - c(A,a)y)x - \eta(a)zx - \eta(a)\psi(a,A)ux \\ \frac{dy}{dt} &= (b(A) - d(A) - c(A,A)y - c(a,A)x)y - \eta(A)zy + \eta(a)\psi(a,A)ux \\ \frac{dz}{dt} &= (1-\gamma)V(\eta(a)x + \eta(A)y)z - (d_v + \eta(a)x + \eta(A)y)z \\ \frac{du}{dt} &= \gamma V \eta(A)yz - (d_v + \eta(a)x + \eta(A)y)u \end{cases} \tag{4}$$

The death rate is taken as $d(a, X) := d(a) + c(a,a)x + c(A,a)y$ where c measures the intensity of competition between individual cells. We also define the transduction probability as

$$\psi(a, A) := \psi_0 \exp\left(-\left(\frac{(a-A)}{\sigma_\psi}\right)^2\right), \tag{5}$$

meaning that the closer the phenotypes are, the more likely the transduction is, with maximal transduction probability being equal to ψ_0.

In a monomorphic population (trait value a), transduction alone has no effect on bacteria's phenotype, hence the phage-bacteria population equilibrium:

$$\begin{cases} \bar{x}(a) &= \frac{d_v}{\eta(a)((1-\gamma)V-1)} \\ \bar{z}(a) &= \frac{1}{\eta(a)}\left(r(a) - \frac{d_v}{\eta(a)((1-\gamma)V-1)}c(a,a)\right) \end{cases} \tag{6}$$

In our numerical simulations, the parameter ranges (see **Supplementary Table 1**) are chosen so that the ecological equilibrium point between a homogeneous bacteria population and a population of viruses always exists: we observe no cycles and no virus or bacteria extinction for the chosen range of traits.

2.3 Mutant invasion

We place ourselves in the case of small mutation steps occurring at an evolutionary timescale, not an ecological one. These two hypotheses allow us to put ourselves in the frame of work described in Geritz et al. 2002 and Geritz 2005. Though there are some mathematical differences induced by our modelling of transduction, transduction rates are typically low enough to make the framework compatible outside of the neighborhood of the evolutionary equilibrium (see **Supplementary note S6** for more details). We can then derive the *invasion implies fixation* result from these papers. This means that no two populations of resident and mutant

individuals phenotypically close are able to coexist except eventually near the evolutionary stable strategy. For one to thrive, the other must go extinct.

We can then divide the evolution of our population in two distinct phases:

1. **Monomorphic phase:** a single type of bacteria population is present along with viruses

2. **Invasion phase:** a single mutant appears in the resident population. The outcome of this phase is either the mutant successfully invades the resident population, driving it to extinction, or it goes extinct itself. Either way, at the end of this phase, only one type of bacteria is present in the system.

We then focus on a single invasion event and derive invasion fitness and invasion probability.

From the previous dynamical system we derive analytically the invasion fitness of mutant trait A on resident trait a. To infer probability and characteristic time of invasion of the mutant population, we come back to the individual scale and couple the process to simple birth and death models.

2.4 Trait substitution sequence and canonical equation

Results from a single invasion event allows us to change the timescale from an ecological perspective to an evolutionary one, resulting in invasion events being resolved instantaneously. If the mutation rate is at the same order of magnitude, each invasion event will resolve into a 'winner' and a 'loser', meaning at each time only one trait will be present in the population if we start off with only one trait. This frame of reference allows us to define the trait substitution sequence (Metz et al. 1995, Champagnat 2006) and the canonical equation (Dieckmann and Law 1996, Champagnat, Ferrière, et al. 2002). A full analysis on this matter can be found in the **Supplementary note S4**.

We rescale time by $K\mu$, K population size and μ mutation rate. At this time scale, we can then see the evolution of our trait as a jump process, the trait substitution process, which can jump from trait a to trait A with rate

$$p(A, a) \times \bar{x}(a)b(a)m(a, dA) \tag{7}$$

where p is the probability of invasion of a mutant trait A on a resident population of trait a, $\bar{x}(a)$ is the equilibrium point of the resident population of trait a and $m(a, dA)$ is the mutation kernel from trait a to trait A, which we can write $m(a, dA) = m(A - a)dA$. This stochastic process is the *Trait Substitution Sequence* (TSS).

Taking the limit of arbitrarily small mutations happening fast (relative to this timescale), the trait substitution sequence converges towards the canonical equation (Champagnat, Ferrière, et al. 2002,Champagnat and Méléard 2011). This canonical equation holds a stochastic part when transduction occurs, which will be a key point of our study.

2.5 Simulations

In order to test our different analytic results, we implemented a rigorous Gillespie algorithm modeling populations and trait distribution at an individual level. We adapted the algorithm descibed in Champagnat 2006 to fit our model. We also simulated the TSS in order to derive adaptation speed and compare the results to our analytical ones.

Default values for the parameters can be found in **Supplementary Table 1**. We can note here that birth, death and extinction rates are taken as constant, respectively equal to b_0, d_0, η_0.

3 Results

3.1 Dynamics of bacterial adaptation without transduction

First we consider a baseline scenario in which the evolving bacterial population is exposed to viral infection but no transduction occurs, $\psi_0 = 0$. Here the invasion fitness $F(A, a)$ of a mutant trait A in a resident population of trait a is

$$F(A, a) = r(A) - c(a, A)\bar{x}(a) - \eta(A)\bar{z}(a). \tag{8}$$

The term $r(A) - c(a, A)\bar{x}(a)$ is the invasion fitness in the absence of viral infection. The additional term $-\eta(A)\bar{z}(a)$ measures the negative effect of infection on bacterial fitness. As expected, this negative effect is stronger if the equilibrium viral population, $\bar{z}(a)$, or the

infection rate, $\eta(A)$, is larger. From equation (8) it follows that the invasion probability is $p(A,a) = [F(A,a)]_+/b(A)$, with a characteristic time of invasion of order $O(\log(K))$, K being the order of magnitude of the size of the population.

Assuming mutations of small phenotypic effects, the dynamics of the mean trait value obey the so-called "canonical equation of adaptive dynamics"

$$\mathrm{d}a_t = \sigma_0^2(a_t)\bar{x}(a_t)\partial_1 F(a_t, a_t) \tag{9}$$

where $\sigma_0(a)$ is the standard deviation of the mutation kernel. The direction of adaptation is given by the selection gradient $\partial_1 F(a_t, a_t)$, while at any time t the rate of adaptation is a function of the selection gradient, population size, and mutation size at that time. For standard choices of functions r, c and η, there is a single "evolutionarily singular" trait value, a^*, at which the selection gradient vanishes. If selection is directional, the trait distribution evolves towards a^*. If selection is stabilizing, the trait distribution remains in a neighborhood of a^* (a numerical example is shown in figure 4 a-c.)

Selection being stabilizing vs. disruptive around the singularity a^* is predicted by the criterion for evolutionary branching. For the model specifications used in numerical simulations, there is no branching, hence stabilizing selection, for $\sigma_K < \sigma_C$, and branching, hence disruptive selection, for $\sigma_K > \sigma_C$ (figure 3a).

3.2 Mutant invasion with transduction

With transduction ($\psi_0 > 0$), invasion fitness $S(A,a)$ of a mutant trait value A in a resident population of trait a is given by

$$S(A,a) = F(A,a) + \frac{\gamma}{1-\gamma}\eta(A)\bar{z}(a)\psi(a,A). \tag{10}$$

Transduction alters invasion fitness by adding the positive term $\frac{\gamma}{1-\gamma}\eta(A)\bar{z}(a)\psi(a,A)$, hence $S(A,a) > F(A,a)$. Thus, invasion fitness with transduction is always greater than invasion fitness without transduction. As a consequence, mutants that are selected against (negative invasion fitness F) in the absence of transduction may be favored and invade with transduction ($S > 0$). This is always the case for negatively selected mutants ($F < 0$) that are

phenotypically close to the resident phenotype. More generally, with transduction, any mutant of sufficiently small effect is predicted to invade.

Transduction also affects the mutant-resident dynamics following mutant invasion. The potential for "back invasion" by the resident population once rare is determined by the "back invasion" fitness (see **Supplementary Note S2**)

$$\tilde{S}(a, A) \;=\; F(a, A) - \frac{\gamma}{1-\gamma}\eta(a)\bar{z}(A)\psi(a, A) \tag{11}$$

By making a negative contribution to back invasion fitness, transduction generally hampers back invasion. In particular, for any mutation of small effect, transduction always drives invasion ($S(A, a) \approx \frac{\gamma}{1-\gamma}\eta(A)\bar{z}(a)\psi(a, A) > 0$) and always prevents back invasion ($\tilde{S}(a, A) \approx -\frac{\gamma}{1-\gamma}\eta(a)\bar{z}(A)\psi(a, A) < 0$).

In Supplementary note S3 we show that transduction always increases the probability of invasion of a mutant. We also whow that for small transduction rates (that is the case when mutations have small phenotypic effects) and large viral burst sizes ($V >> 1$, which is typically the case), the probability of invasion with transduction is

$$p(A, a) \approx [S(A, a)]_+/b(A). \tag{12}$$

The characteristic time of invasion stays of order $\log K$ where K quantifies the order of magnitude of population size. In the next section, we use equation (12) to derive the rate of adaptation.

Of note, equations (10) and (12) imply $S(a, a) > 0$ and $p(a, a) > 0$ for any trait value a. Thus, a mutant that is phenotypically identical to the resident type has positive invasion fitness, as opposed to zero invasion fitness in models without transduction. This is due to our explicit genotype-phenotype map: albeit phenotypically identical to the resident type, the mutant type is genetically different (by definition, it carries a new mutation) and our model of transduction accounts for this. Using notations introduced earlier, we have $\mathcal{M}_{res} = \emptyset$ and $\mathcal{M}_{mut} = \{1\}$, and our model predicts $S(\mathcal{M}_{mut}, \mathcal{M}_{res}) > 0$. However, $S(\mathcal{M}_{res}, \mathcal{M}_{res}) = 0$, which is consistent with the expectation that a strain that is virtually identical, both genetically and phenotypically, should have zero invasion fitness.

Figure 2c shows the invasion probability for mutants phenotypically identical to their

resident progenitor, $p(a, a)$, which also provides an approximation for the invasion probability of mutations of small phenotypic effects. The probability $p(a, a)$ is higher near the evolutionary singularity a^*. Thus, due to transduction, the invasion probability of any mutation of small effect increases as the adaptation process brings the population closer to the evolutionary singularity.

3.3 Rate of bacterial adaptation with transduction

Here we assume that mutations have small phenotypic effects. By rescaling time appropriately, we derive a macroscopic model for the dynamics of adaptation which takes the form of a stochastic differential equation or integro-differential equation, depending on the shape of the distribution of phenotypic mutational effects, or "mutation kernel". In **Supplementary Note S4**, we show that the models, hence the dynamics of adaptation that they capture, critically depend on whether the mutation kernel is symmetrical around zero (unbiased mutational effects) or not (biased mutational effects).

3.3.1 Case of unbiased mutation

With unbiased mutation the rate of adaptation is controlled by

$$da_t = u(a_t)\ \partial_1 p(a_t, a_t)\ dt + \sqrt{u(a_t)p(a_t, a_t)}\ dB_t \tag{13}$$

where

$$u(a_t) = \sigma_0^2(a_t) b(a_t) \bar{x}(a_t),$$

$\sigma_0(a)$ is the standard deviation of the mutation kernel when the trait value is equal to a, and B_t is a Brownian motion. The Brownian component arises because transduction causes $p(a, a) > 0$. If there is an evolutionary singularity a^* where $\partial_1 p(a^*, a^*)$ vanishes, then equation 13 shows that the dynamics of adaptation become more stochastic as the population evolves closer to the singularity a^*, because $p(a, a)$ is larger around a^*.

If we assume that the infection rate is independent of the evolving trait a, then the deterministic part of the rate of adaptation with transduction and the rate of adaptation without

transduction are equal. However, numerical simulations of the trait substitution sequence (TSS) show that the actual rate of adaptation with transduction differs from the adaptation rate without transduction, due to an interaction between the deterministic and stochastic drivers of the adaptation dynamics.

We quantify this effect with respect to two key parameters of the bacteria-phage interaction, the rate of infection η and the probability of transduction ψ. A index of adaptation stochasticity, I, can be defined as the ratio between the speed of evolution without transduction and the mean sojourn time of the population in a small neighborhood around the evolutionary singularity (see **Supplementary note S5** for more detail). $I \ll 1$ means that the stochastic component of adaptation drives the process away from evolutionary equilibrium faster than the deterministic driver can bring it back. In this case, the population remains in a maladapted state, i.e. away from the predicted adaptation. Conversely, $I \gg 1$ means that the deterministic driver of adaptation brings the population near its evolutionary equilibrium point faster than the stochastic component upsets adaptation. In this case, adaptation trajectories can be approximated with exponential functions of time t of the form $C\exp(-\alpha t)$. The parameter α then provides a measure of convergence speed that can be compared across models with and without transduction.

Figure 4d shows that for values of ψ_0 and η_0 that are lower that 0.05, $I \approx 1$ and therefore even though convergence towards an evolutionary singularity is expected, the adaptation trajectories strongly fluctuate around the singularity without ever settling. Above these extremely low values and across a broad range of infection parameters ψ_0 and η_0, adaptation speed v increases with higher transduction probability or infection rate (figure 4e,f). Thus, with unbiased mutation, transduction significantly accelerates adaptation, provided the rates of infection or transduction are not too low.

3.3.2 Case of biased mutation

With biased mutation, the mean mutational effect is nonzero and given by

$$\mu_0(a) = \int_{\mathcal{D}} h \; m(a,h) dh$$

Then the characteristic timescale over which adaptation proceeds is set by t/ϵ, where t is the baseline timescale over which the cells' life histories unfold. The t/ϵ timescale is much shorter

than the t/ϵ^2 timescale over which adaptation proceed with unbiased mutation, i.e. adaptation occurs much more rapidly with biased mutation. On this faster timescale, the stochasticity of the adaptation process with transduction is smoothed out, and the adaptation dynamics are governed by a deterministic ordinary equation

$$da_t = \frac{\gamma}{1-\gamma}\psi_0 \ \mu_0(a_t) \ \eta(a_t)\bar{x}(a_t)\bar{z}(a_t) \ dt. \tag{14}$$

In this case, the long-term adaptation dynamics takes steps of the order of the average mutational bias, measured by $\mu_0(a)$, at a rate given by the population rate of transduction events, $\frac{\gamma}{1-\gamma}\psi_0\eta(a_t)\bar{x}(a_t)\bar{z}(a_t)$.

3.4 Effect of transduction on bacterial adaptive diversification

Without transduction, the model predicts conditions under which evolutionary branching is expected around the evolutionary singularity a^* (figure 3a). Mathematically describing this evolutionary process with transduction goes beyond the scope of this paper. A large set of numerical simulations were produced instead. The simulations consistently show that even the lowest level of transduction widens the trait range over which a mutant can invade in a neighborhood of the evolutionary singularity (figure 3). None-the-less, evolutionary branching occurs under the conditions predicted in the absence of transduction, and the fact that back invasion fitness is always negative does not seem to slow the process.

Below a threshold on the transduction rate (low ψ_0), a third, intermediate branch can evolve and station around the evolutionary singularity, which is a fitness minimum (figure 3b-d). Competition exerted by individuals in the intermediate branch then select for further divergence of the outer branches (figure 3d). Thus, transduction tends to increase the genetic and phenotypic diversity and phenotypic range that evolves in the population. Above the threshold transduction rate, large stochastic variations manifest in the outer two branches, which then tend to outcompete any middle branch (figure 3e-f). Also, large stochastic excursions of one branch away from the evolutionary singularity may allow the other branch to stay closer to the singularity (figure 3e-f). Thus, transduction makes it possible for a subpopulation to station around a fitness minimum. Other example simulations are shown in Supplementary figure 1.

Discussion

Transduction is known to have a great deal of importance in shaping ecosystems (Jiang and Paul 1998), but its effects can both favor and hinder bacterial adaptation. To resolve these conflicting effects, we extended a trait-based approach to include a simple genotype-phenotype map in a trait-based ecological model of bacteria-phage population dynamics. By doing so, we were able to capture the phenotypic effect of genetic mobility driven by transduction. Our hybrid – genetic and trait-based – model uses an infinite-site model approach while neglecting gene dosage effects. As a consequence, the transduction of a mutant allele to a different genotype (wild-type or other mutants) will alter the latter's phenotype, whereas the transduction of the wild-type allele to any mutant genotype will not change their phenotype.

Though multiple studies have been conducted regarding global impact of horizontal gene transfer (Nazarian et al. 2018 for example), few do it with a mechanistic perspective. Compartmental models are used (as in Niehus et al. 2015), and gene flow is broadly described as rates of transfer from one compartment to another. Delving deeper into each transfer mechanism is interesting to better understand and predict the consequences of each type. Individual-centered mechanistic models have been developed for transformation (Mao and Lu 2016) and conjugation (Billiard et al. 2016), but to the best of our knowledge such a work has not been undertaken for transduction. That is why we propose to model the influence of viruses of adaptation in our study.

The most striking modelling decision that we made was to study the nonhomologous end-joining transduction mechanism (Popa et al. 2011). This decision has the direct consequence that horizontal gene transfer now relies on genotype and not only phenotype. This approach departs from classical *adaptive dynamics* approach to borrow methods from *population genetics*, and was chosen in order to stick as closely as possible to the biological reality underlying nonhomologous end joining transduction. Since transduction is done through integrating foreign genes inside a host genome, we can't rely on phenotype alone (therefore, on trait alone) to describe such an event. Our modelling allows us to track the genealogy and the individual mutations that occurred, as it is done in population genetics, while maintaining classical adaptive dynamics results for all matters not involving transduction. We believe this novel way of modelling gene integration could be the foundation for further studies involving the

need for both population genetics and adaptive dynamics.

The goal of our model was to study the impact of transduction on the incremental adaptation of bacteria as opposed to the transfer of whole ecological functions as the result of horizontal gene transfer.

Our first step was to quantify the effect viruses can have on adaptation in a model where transduction didn't occur ($\psi_0 = 0$). The main difference brought by viruses was the switch between a 'bottom-up' limitation of bacterial growth due to nutrients to a 'top-down' limitation set by viruses (consistent with Sieradzki et al. 2019). This can be seen in (6), which shows bacteria equilibrium relying solely on viral parameters. In addition to that, we see that the invasion fitness is that of a classical monomorphic system in which viruses add a supplementary mortality term $\eta(a)\bar{z}(a)$ which slows down adaptation.

We found that transduction can upset the direction of natural selection in an invasion event. We see in the expression (10) that we can divide the invasion fitness S into two terms: the invasion fitness in absence of transduction F and a 'transfer term'. The presence of a transfer term for an invasion fitness involving horizontal gene transfer is not unheard of (see Billiard et al. 2016 for the case of plasmids). We can divide this transfer term into three factors: $\eta(A)\bar{z}(a)$, $\frac{\gamma}{1-\gamma}$, and $\psi(A, a)$. The first factor represents the rate at which mutant bacteria will interact with viruses present at resident ecological equilibrium, each of these interaction being a potential opportunity for releasing GTAs. The second factor represents the relative abundance of GTAs released by a mutant infection in regards to actual viruses. The two factors combined represent the strength of potential transduction events compared to infection events, which hinder a successful invasion. Finally, the last term represent the proportion of those potential transduction events that come to fruition. The three factors combined thus indicate the strength of transduction, and prove that transduction can upset selection, as it can change a naturally deleterious trait A (i.e. $F(A, a) < 0$) into an advantageous one ($S(A, a) > 0$).

Helping fixate a deleterious trait is sometimes called the 'bad gene effect', and it is a well studied fact that horizontal gene transfer can favor it (Moradigaravand and Engelstädter 2014). But one fundamental difference here is that we do not witness a 'reversal of selection' (as in Billiard et al. 2016 for example), but a widening of selection outcomes. We see that the transfer term is positive for any mutant trait that would appear in the resident population.

This means that naturally deleterious traits can indeed become advantageous, but it also means that already advantageous traits become all the more advantageous. This widening can be of great importance, as we see in Figure 2b. This goes to show that transduction helps any mutant invasion, regardless of the 'natural' (i.e. without transduction) direction of selection.

This widening of potential invaders is the highest near the evolutionary equilibrium a^*. Indeed, we see on Figure 2c that the invasion probability is greater near a^*. This has a fairly straightforward ecological explanation. The 'top-down' limitation induced by viruses ensures that approaching the evolutionary equilibrium favors the growth of the viral population instead of bacteria. The closer we get to the evolutionary stable strategy, the more viruses roam free in the environment for an equal number of bacteria, and so the more transduction events take place, favoring the invasion of close phenotypical neighbors. This allows for a constant turn-over of phenotypes, all close to the evolutionary stable strategy.

Therefore, transduction is at the crux of trait evolution stochasticity. The stochastic form that takes the canonical equation (13) is a direct consequence of the widening of invasion outcomes. Indeed, since small phenotypical variations on both sides of resident trait have a positive probability of invading, there is no longer only one evolutionary direction. Even though one direction is favored (as can be seen by the drift factor of the canonical equation), random perturbations can occur on either side of resident trait. Perturbations are highest near the evolutionary equilibrium, assuring that traits far enough from equilibrium will tend to converge towards a neighborhood of an evolutionary stable strategy.

The uncertainty on invasion outcome directly reflects on the evolutionary process, giving it its stochastic nature. The effect of transduction on the canonical equation (13) can be compared to the effect of genetic drift. Indeed, we see in Champagnat, Lambert, et al. 2007 that both canonical equations share a stochastic part proportional to the square root of the invasion probability of a trait over itself. In our case, the fact that this term is nonzero comes from the effect of transduction, whereas in Champagnat, Lambert, et al. 2007, it comes from the fluctuations in population size. The effect is inherently the same, providing the possibility for naturally deleterious traits to invade a more adapted resident population while keeping the invasion probability of an advantageous trait positive.

A natural trade-off then comes to mind: on one hand, by boosting the probability of invasion for a mutant that is closer to the evolutionary stable strategy, transduction contributes to accelerating adaptation. On the other hand, transduction can also knock adaptation back by allowing a naturally deleterious mutant to invade the resident population. We find that at low rates, transduction helps adaptation. Indeed, we see on Figure 4 that adaptation speed consistently increases with transduction, but the Supplementary figure 2 shows that at higher rates, the gain in speed is counteracted by a greater stochasticity: the process is no faster than the values shown on Figure 4, but the stochasticity is much higher, meaning the process won't stay long close to the optimal value.

The fact that trait fluctuations still occur near the evolutionary stable strategy implies that the bacterial population will never be fully adapted to a static environment: by favoring mutant fixations, transduction lowers the mean fitness of the population. This confirms that transduction doesn't confer a selection advantage in static environments, but that the increased adaptation speed is beneficial in changing environments (Raz and Tannenbaum 2010).

The observed acceleration is consistent with experimental data showing that laterally transferred genes usually undergo faster evolution (Hao and Golding 2006, Marri et al. 2007) and that horizontal gene transfer helps adaptation in a microevolutionary way (Baltrus et al. 2008, Cohen et al. 2006). These studies show how mutations appear more frequently for laterally transferred genes: this can be seen in the TSS (Figure 4), where stochastic fluctuations of trait happen at a timescale one order of magnitude faster than general adaptation. In a proper sense, the mutations do not 'appear' more often, but they tend to fixate in the population more easily.

Adaptation acceleration could have numerous implications. Transduction is supected to have an important role in shaping genetic materials in the oceans: we estimate as much as 10^{14} transduction events a year in Tampa Bay, Gulf of Mexico alone (Jiang and Paul 1998). In a context of rapidly changing environments because of climate change, a better understanding of bacterial response to transduction is a step towards integrating this important mechanism in global biogeochemical models of the oceans. Further studies should seek to integrate an eco-evolutionary response from viruses themselves to better integrate the multiple coevolutionary

responses we find in natural environments, such as 'kill the winner' or 'arms race' mechanisms.

Once settled in the neighborhood of the evolutionary equilibrium, transduction can still upset the phenotypical landscape we would expect. We looked at the process under regular branching conditions (Dieckmann and Doebeli 1999) to see the effect of transduction on phenotypical diversity.

We see on the individual-based models (Supplementary figure 1) that two major effects of transduction can explain the different patterns: genome mixing and trait stochasticity. When branching occurs, transduction drives the creation of one branch between the two 'natural' ones by mixing their genome (as is explained in Figure 1b). This branch follows either one of the natural branches and is parallel to it: this is due to the fact that they only survive thanks to horizontal gene transfer, and would go extinct without the flux of individual coming from the original branch. This effect can be seen when transduction rates are low enough, reducing the importance of trait stochasticity. Conversly, when transduction rates are high, the stochasticity of trait evolution can upset the balance that made viable a third branch. Indeed, when branching occurs, if either one of the branches stays close to the evolutionary equilibrium due to the stochasticity induced by transduction, the competition it exerts on the potential third branch prevents it from even appearing (Figure 3e,3f). This means that at low enough rates, transduction can help phenotypical diversity through gene-mixing, but that at high rates the induced stochasticity renders the resulting branches unstable, therefore hindering phenotypical diversity. Our study also unravels a novel mechanistic explanation to why horizontal gene transfer is sometimes thought to help phenotypical diversity when other times it hinders it (Niehus et al. 2015).

In conclusion, this study proves through a mix of analytical and numerical analysis how free-floating viruses contribute to the genetic pool in oceans through the transduction mechanism by accelerating adaptation and adding to the genetic mixing. In a context of booming studies regarding the influence of viruses in global microbial ecology, we believe this adds to the mechanistic understanding of host-virus interactions.

References

[1] Cheryl P Andam, Sarah M Carver, and Sean T Berthrong. "Horizontal gene flow in managed ecosystems". In: *Annual Review of Ecology, Evolution, and Systematics* 46 (2015), pp. 121–143.

[2] David A Baltrus, Karen Guillemin, and Patrick C Phillips. "Natural transformation increases the rate of adaptation in the human pathogen Helicobacter pylori". In: *Evolution: International Journal of Organic Evolution* 62.1 (2008), pp. 39–49.

[3] Sylvain Billiard et al. "The effect of competition and horizontal trait inheritance on invasion, fixation, and polymorphism". In: *Journal of theoretical biology* 411 (2016), pp. 48–58.

[4] Mya Breitbart et al. "Phage puppet masters of the marine microbial realm". In: *Nature Microbiology* 3.7 (2018), p. 754.

[5] Nicolas Champagnat. "A microscopic interpretation for adaptive dynamics trait substitution sequence models". In: *Stochastic processes and their applications* (2006).

[6] Nicolas Champagnat, R Ferrière, and G Ben Arous. "The canonical equation of adaptive dynamics: a mathematical view". In: *Selection* 2.1-2 (2002), pp. 73–83.

[7] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. "Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models". In: *Theoretical Population Biology* (2006).

[8] Nicolas Champagnat and Sylvie Méléard. "Polymorphic evolution sequence and evolutionary branching". In: *Probability Theory and Related Fields* 151.1-2 (2011), pp. 45–94.

[9] Nicolas Champagnat, Amaury Lambert, et al. "Evolution of discrete populations and the canonical diffusion of adaptive dynamics". In: *The Annals of Applied Probability* 17.1 (2007), pp. 102–155.

[10] Elisheva Cohen, David A Kessler, and Herbert Levine. "Analytic approach to the evolutionary effects of genetic exchange". In: *Physical Review E* 73.1 (2006), p. 016113.

[11] Roberto Danovaro et al. "Marine viruses and global climate change". In: *FEMS microbiology reviews* 35.6 (2011), pp. 993–1034.

[12] Fernando De la Cruz and Julian Davies. "Horizontal gene transfer and the origin of species: lessons from bacteria". In: *Trends in microbiology* 8.3 (2000), pp. 128–133.

[13] Ulf Dieckmann and Michael Doebeli. "On the origin of species by sympatric speciation". In: *Nature* 400.6742 (1999), p. 354.

[14] Ulf Dieckmann and Richard Law. "The dynamical theory of coevolution: a derivation from stochastic ecological processes". In: *Journal of mathematical biology* 34.5-6 (1996), pp. 579–612.

[15] Kevin R Foster et al. "The evolution of the host microbiome as an ecosystem on a leash". In: *Nature* 548.7665 (2017), p. 43.

[16] Stefan AH Geritz. "Resident-invader dynamics and the coexistence of similar strategies". In: *Journal of Mathematical Biology* 50.1 (2005), pp. 67–82.

[17] Stefan AH Geritz et al. "Invasion dynamics and attractor inheritance". In: *Journal of Mathematical Biology* 44.6 (2002), pp. 548–560.

[18] James PJ Hall, Michael A Brockhurst, and Ellie Harrison. "Sampling the mobile gene pool: innovation via horizontal gene transfer in bacteria". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1735 (2017), p. 20160424.

[19] Weilong Hao and G Brian Golding. "The fate of laterally transferred genes: life in the fast lane to adaptation or death". In: *Genome Research* 16.5 (2006), pp. 636–643.

[20] Sunny C Jiang and John H Paul. "Gene transfer by transduction in the marine environment". In: *Applied and environmental microbiology* (1998).

[21] Amaury Lambert et al. "The branching process with logistic growth". In: *The Annals of Applied Probability* 15.2 (2005), pp. 1506–1535.

[22] Jeffrey G Lawrence. "Gene transfer in bacteria: speciation without species?" In: *Theoretical population biology* 61.4 (2002), pp. 449–460.

[23] Junwen Mao and Ting Lu. "Population-dynamic modeling of bacterial horizontal gene transfer by natural transformation". In: *Biophysical journal* 110.1 (2016), pp. 258–268.

[24] Pradeep Reddy Marri, Weilong Hao, and G Brian Golding. "The role of laterally transferred genes in adaptive evolution". In: *BMC evolutionary biology* 7.1 (2007), S8.

[25] Johan AJ Metz et al. "Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction". In: (1995).

[26] Adam Monier et al. "Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus". In: *Genome research* 19.8 (2009), pp. 1441–1449.

[27] Danesh Moradigaravand and Jan Engelstädter. "The impact of natural transformation on adaptation in spatially structured bacterial populations". In: *BMC evolutionary biology* 14.1 (2014), p. 141.

[28] Colin B Munn. "Viruses as pathogens of marine organisms—from bacteria to whales". In: *Journal of the Marine Biological Association of the United Kingdom* 86.3 (2006), pp. 453–467.

[29] Phillip Nazarian, Frances Tran, and James Q Boedicker. "Modeling multispecies gene flow dynamics reveals the unique roles of different horizontal gene transfer mechanisms." In: *Frontiers in microbiology* 9 (2018), p. 2978.

[30] Rene Niehus et al. "Migration and horizontal gene transfer divide microbial genomes into multiple niches". In: *Nature communications* 6 (2015), p. 8924.

[31] Howard Ochman, Jeffrey G Lawrence, and Eduardo A Groisman. "Lateral gene transfer and the nature of bacterial innovation". In: *nature* 405.6784 (2000), p. 299.

[32] Ovidiu Popa et al. "Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes". In: *Genome research* 21.4 (2011), pp. 599–609.

[33] Yoav Raz and Emmanuel Tannenbaum. "The influence of horizontal gene transfer on the mean fitness of unicellular populations in static environments". In: *Genetics* 185.1 (2010), pp. 327–337.

[34] Rosemary J Redfield. "Evolution of bacterial transformation: is sex with dead cells ever better than no sex at all?" In: *Genetics* 119.1 (1988), pp. 213–221.

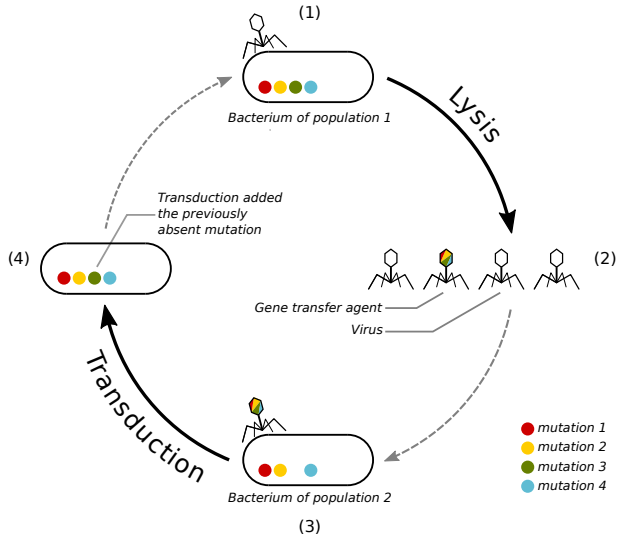
[35] Ella T Sieradzki et al. "Dynamic marine viral infections and major contribution to photosynthetic processes shown by spatiotemporal picoplankton metatranscriptomes". In: *Nature communications* 10.1 (2019), p. 1169.

[36] T Frede Thingstad and Risto Lignell. "Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand". In: *Aquatic Microbial Ecology* (1997).

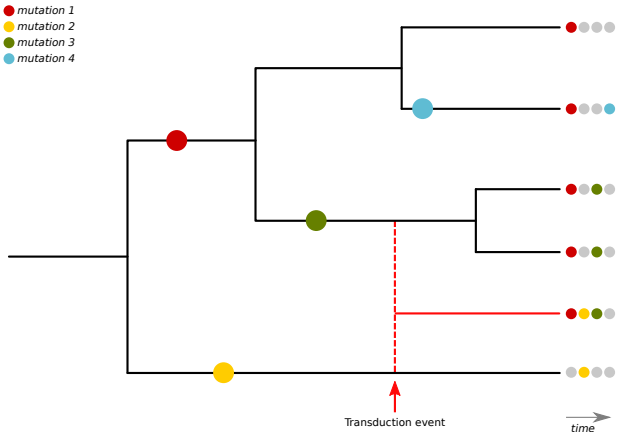
[37] Marie Touchon, Jorge A Moura de Sousa, and Eduardo PC Rocha. "Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer". In: *Current opinion in microbiology* 38 (2017), pp. 66–73.

[38] Erik J Van Hannen et al. "Changes in Bacterial and Eukaryotic Community Structure after Mass Lysis of Filamentous Cyanobacteria Associated with Viruses". In: *Applied and environmental microbiology* (1999).

[39] K Eric Wommack and Rita R Colwell. "Virioplankton: viruses in aquatic ecosystems". In: *Microbiol. Mol. Biol. Rev.* 64.1 (2000), pp. 69–114.



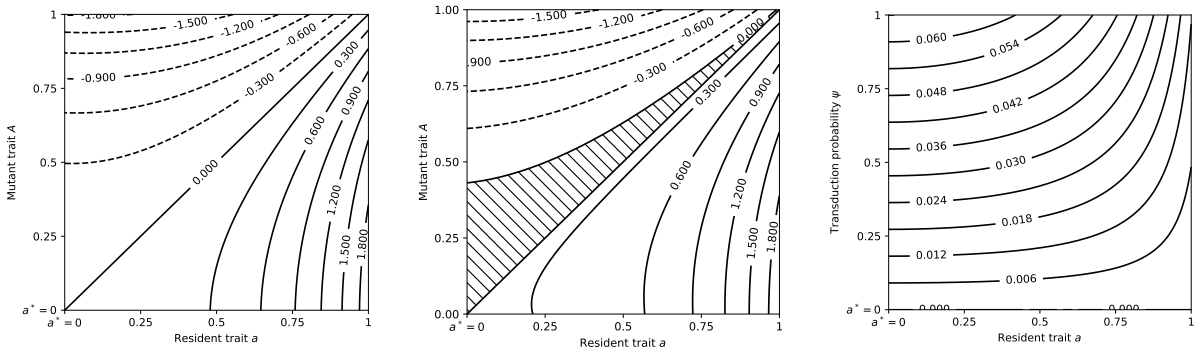
(a)



(b)

Figure 1

Figure 1. The influence of transduction on (a) population dynamics, (b) population genetics. (a) (1) Infection from a virus which results in (2) a viral burst. Some of the viral particles released from the burst are gene transfer agents, which can (3) interact with bacteria through (4) transduction, transferring previously absent genes. **(b)** Black lines represent vertical lineage in which various mutations appear, and the red line represents the resulting lineage after a transduction event. This example shows the genetical state of the population after a transduction event, with the creation of a new genotype comprised of two different genotypes.



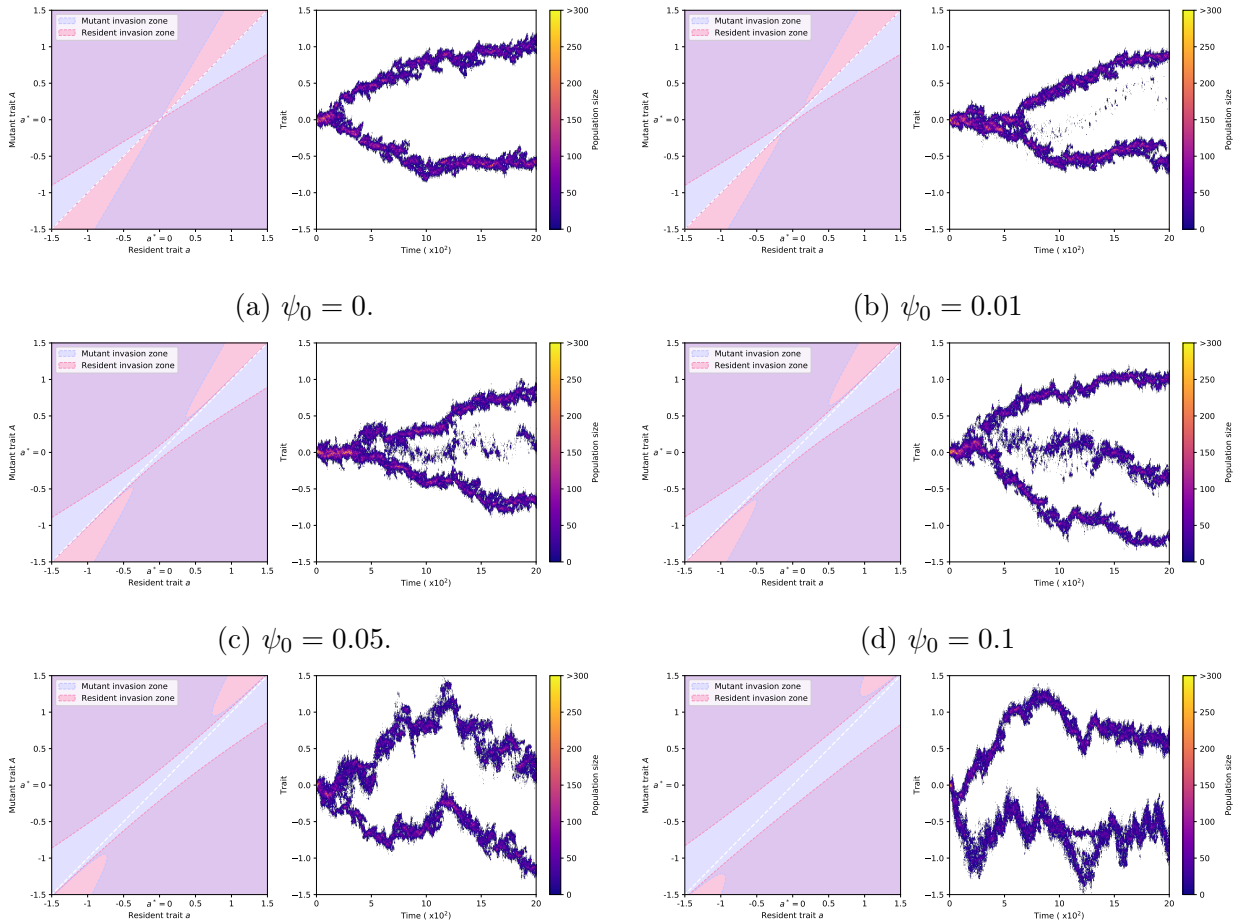
(a) Contour for $F(A, a)$.

(b) Contour for $S(A, a)$.

(c) Contour for $p(a, a)$.

Figure 2

Figure 2. Invasion fitness and probability of invasion on the first bissector. (a) and **(b)** represent contours for functions S and F for $\psi_0 = 1.0$ and $\eta_0 = 1.0\text{E}-5$. Dotted lines represent negative values of the functions, and the hatched zone represents values of a and A for which $F(A, a) < 0$ and $S(A, a) > 0$. **(c)** Invasion probability $p(a, a)$ for different values of trait a and ψ_0. Other parameters are set to their default values (see Supplementary Table 1).



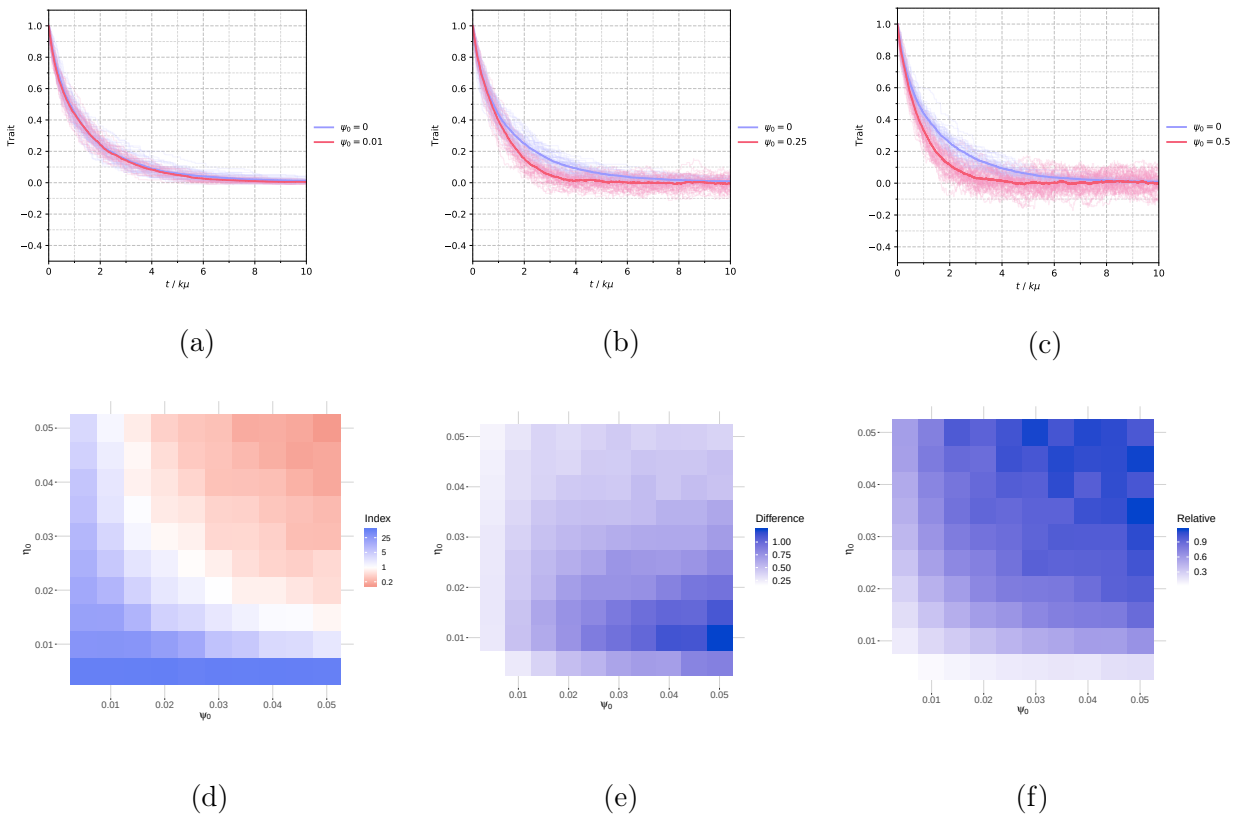
(a) $\psi_0 = 0.$ (b) $\psi_0 = 0.01$

(c) $\psi_0 = 0.05.$ (d) $\psi_0 = 0.1$

(e) $\psi_0 = 0.25$ (f) $\psi_0 = 0.5$

Figure 3

Figure 3. Pairwise invasibility plot for different set of parameters and corresponding trait distribution simulations. For each parameter set, the blue zone (resp. red zone) in the left figure represents traits a, A for which mutant invasion fitness (resp. resident invasion fitness) is positive. Thus, the overlap of the two zones represent possible values of coexistence between mutant ant resident populations. The right figure shows the result of a Gillespie algorithm was used to study branching under transduction. Values of ψ_0 were set depending on the figure, and η_0 was set to $1.0\text{E} - 5$. A change was made to other default parameters, namely $\sigma_K = 2$ and $\sigma_C = 1$ to show the implications of transduction on potential branching. Other parameters are set to their default values (see Supplementary Table 1).



(a) (b) (c)

(d) (e) (f)

Figure 4

Figure 4. Stochasticity and adaptation speed. Trait substitution sequence comparison for $\eta_0 = 1.0\text{E} - 5$ and **(a)** $\psi_0 = 0.01$ **(b)** $\psi_0 = 0.25$ **(c)** $\psi_0 = 0.5$ **(d)** Value of stochasticity index I. Green tiles represent values of $I > 1$ and red tiles represent values of $I < 1$. **(e)** Average absolute difference in adaptation speed α between transduction models and no-transduction models for 100 simulations of each parameter set. **(f)** Average relative difference in adaptation speed α between transduction models and no-transduction models for 100 simulations of each parameter set. Other parameters are set to their default values (see Supplementary Table 1).